Beyond Gaia DR3:
tracing the [α/M] – [M/H] bimodality from the Inner to the outer Milky Way disc with *Gaia* RVS and Convolutional Neural-Networks

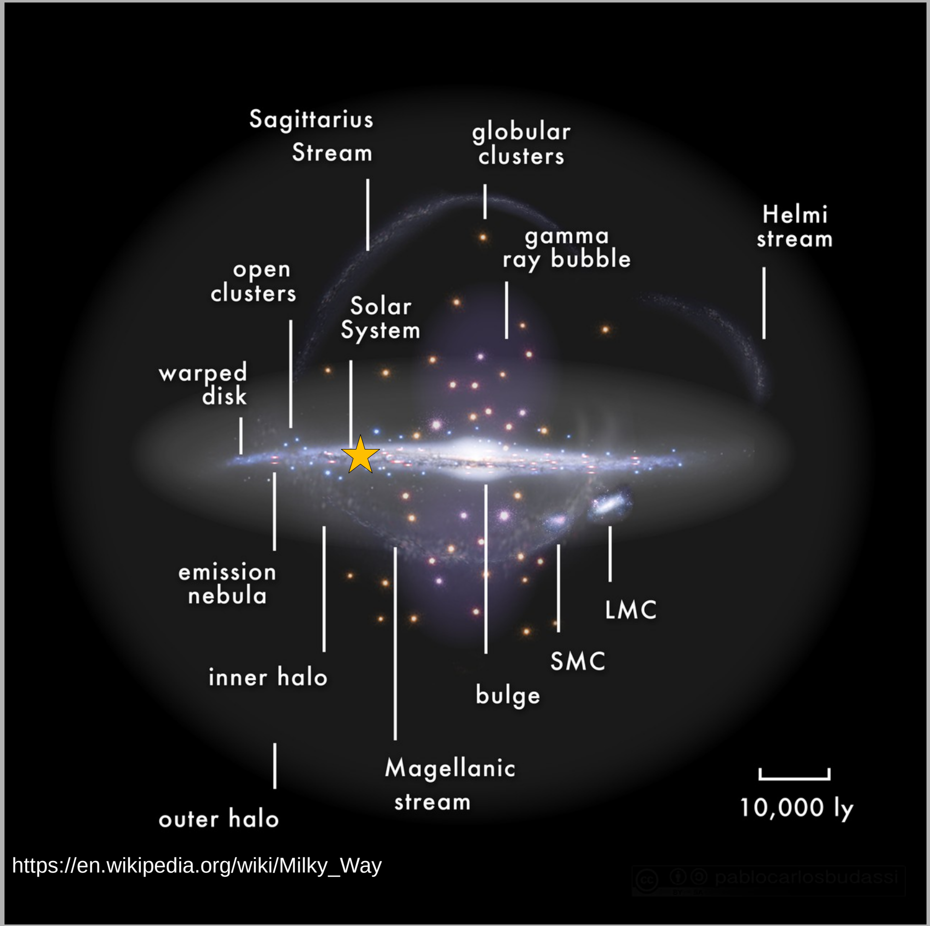
Guillaume Guiglion (GG; LSW / MPIA, guiglion@mpia.de)
& Samir Nepal (AIP)

The Milky Way Revealed by Gaia:
The Next Frontier (Sept. 5-7, 2023)

Galactic Archaeology

→ studying the formation and evolution of the Milky Way and it’s local volume

→ **need for stellar chemistry, kinematics & ages**

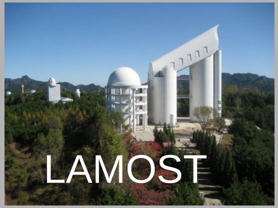


https://en.wikipedia.org/wiki/Milky_Way

RAVE

APOGEE

GaiaESO



GALAH

4MOST

SDSS

gaia

MSE

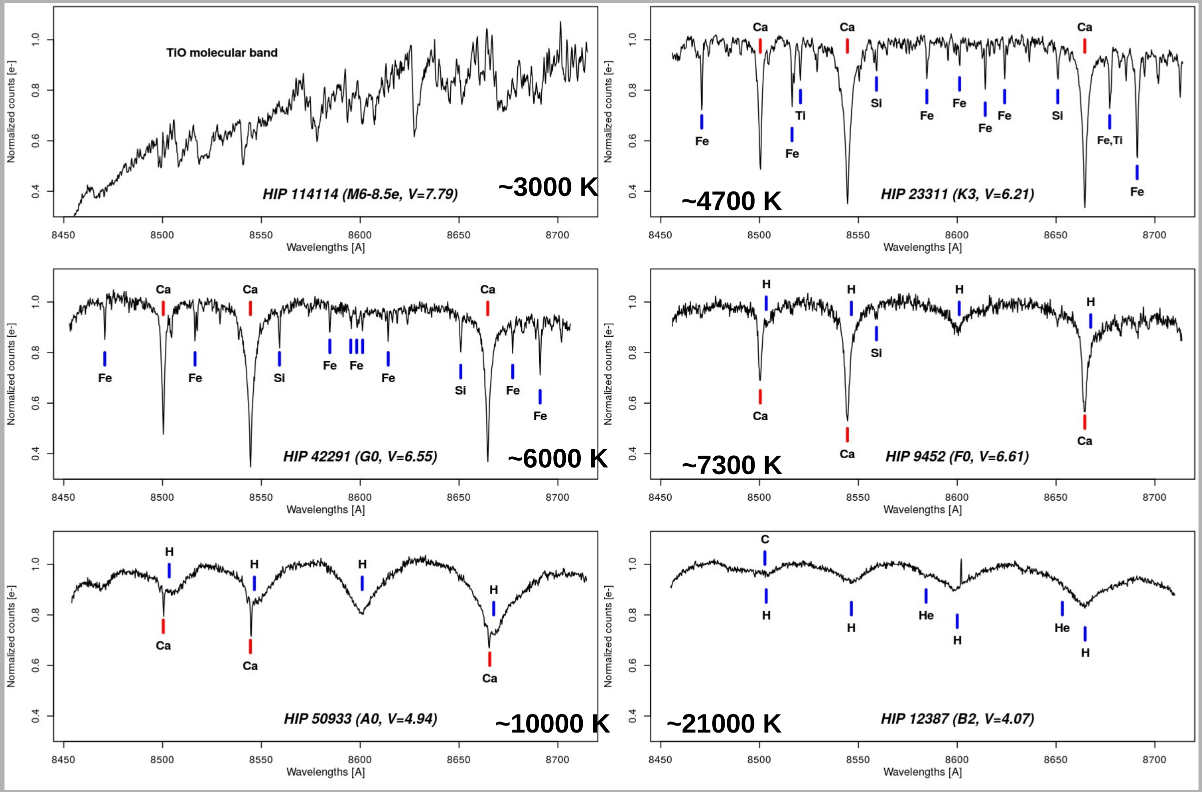
MOONS

Maunakea Spectroscopic Explorer

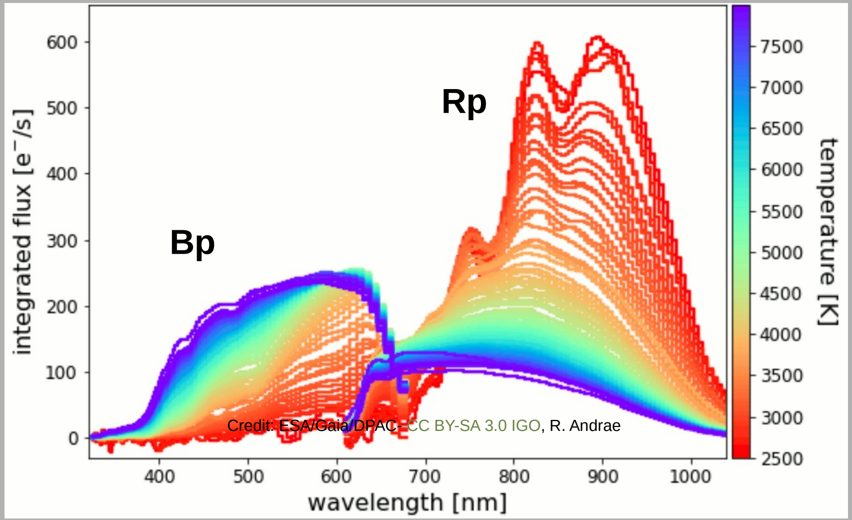
Huge data analysis challenge !!

What Gaia DR3 gave us:

→ 10^6 RVS spectra, R~11500 (Katz et al. 2022)



→ 220 millions BP&RP spectra R~30-100 (De Angeli et al. 2022)



→ 1.5×10^9 parallaxes (Lindegren et al. 2021)

→ 1.8×10^9 G mags

→ 1.5×10^9 BP & RB mags

→ See Recio-Blanco et al. 2023 + talk for standard spectroscopic analysis of RVS spectra

Can we exploit in a homogeneous way
Gaia **spectra (RVS + BP/RP)**
magnitudes (G, Bp, Rp)
and parallaxes
for supercharged stellar
parametrization ?

Analysis of the 1 million *Gaia* RVS-spectra with CNNs

Beyond *Gaia* DR3: tracing the [α/M] − [M/H] bimodality from the Inner to the outer Milky Way disc with *Gaia* RVS and Convolutional Neural-Networks

G. Guiglion[1], S. Nepal[2,3], C. Chiappini[2], S. Khoperskov[2], G. Traven[4], A. B. A. Queiroz[2], M. Steinmetz[2], M. Valentini[2], Y. Fournier[2], A. Vallenari[5], K. Youakim[6], M. Bergemann[1], S. Mészáros[7,8], S. Lucatello[9,10], R. Sordo[5], S. Fabbro[11], I. Minchev[2], G. Tautvaišienė[12], Š. Mikolaitis[12], J. Montalbán[13]

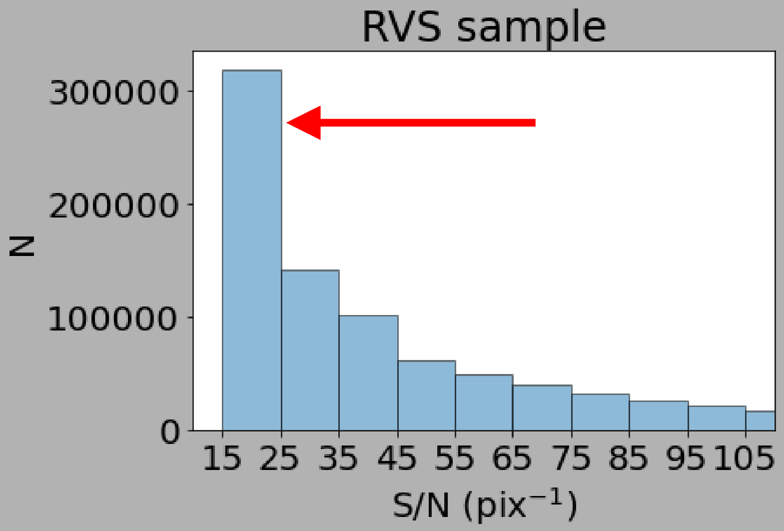
→ **Under revision :)**

Motivations and goals:

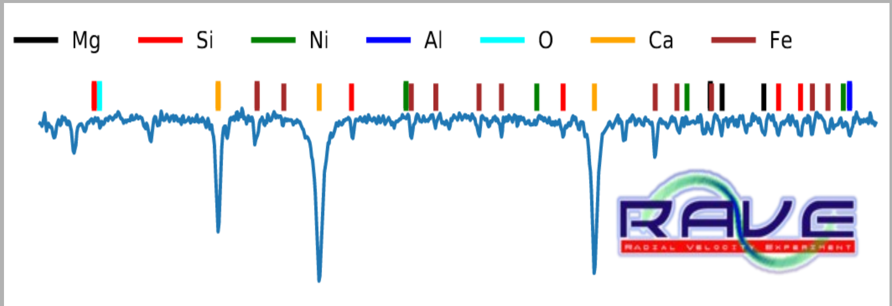
→ **Use homogeneously the full *Gaia* data product**

→ **Leverage the low-S/N RVS sample**
No GSP-Spec labels with 13 “good” flags within 15<S/N<25

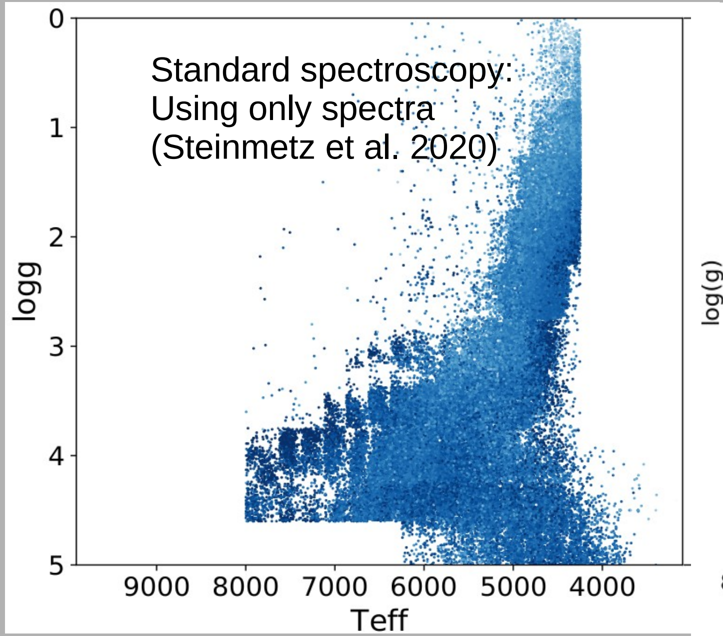
→ **Set the machine-learning path for Gaia data analysis (DR4 in 2025, DR5 in 2027)**

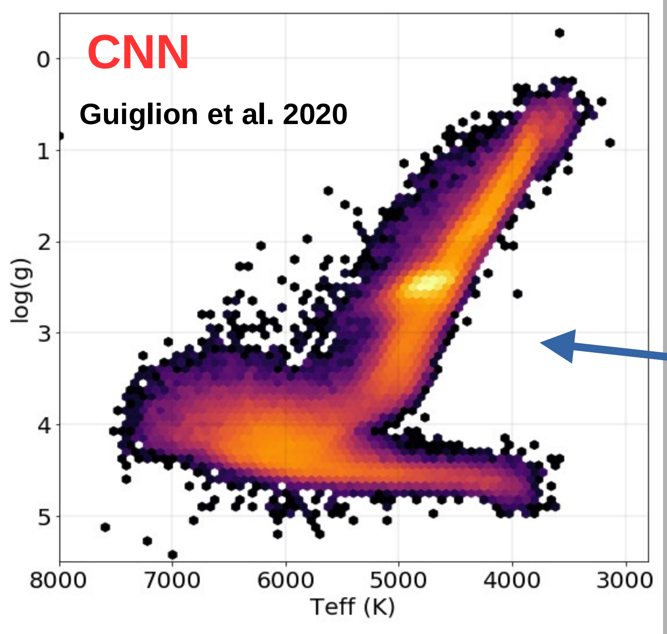


Our experience with CNNs and *Gaia*-like spectra



→ **GG et al. 2020:**
1st application of CNNs combining RAVE spectra, *Gaia* magnitudes, and parallaxes





Breaking the spectral degeneracies

→ **Recent CNN developments in Nepal, GG et al. 2023 and Ambrosch, GG et al. 2023**

Analysis of the 1 million *Gaia* RVS-spectra with CNNs

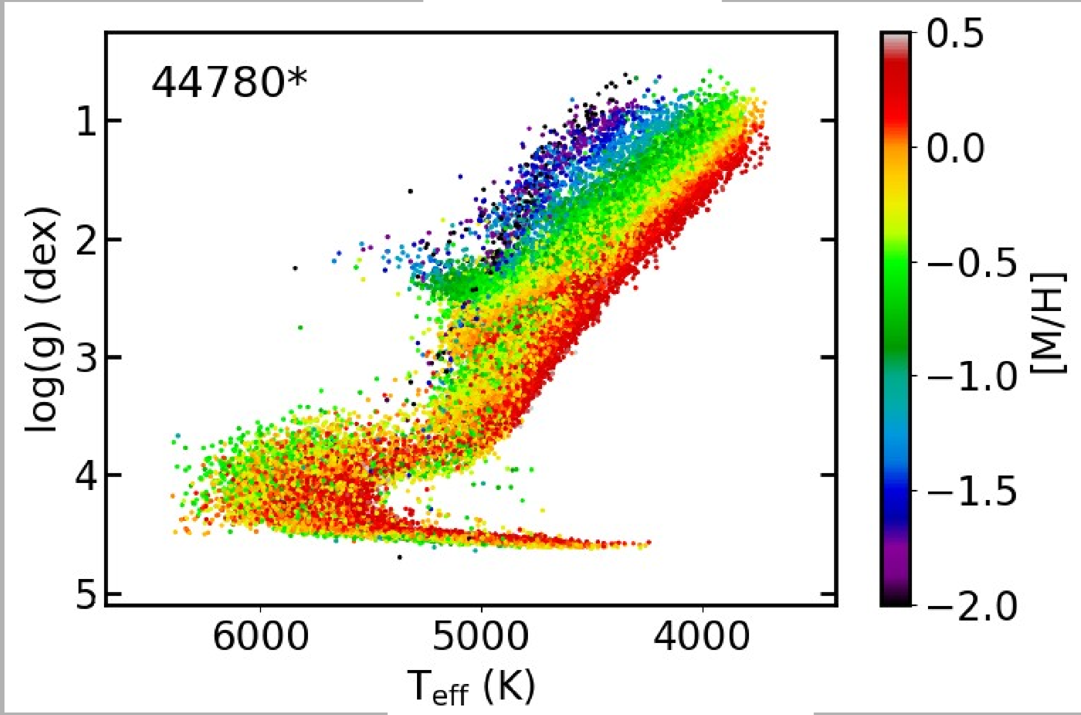
Training sample

R~22000

Knowledge transfer from high-quality high-res APOGEE labels
T_{eff}, log(g), [M/H], [α/M], [Fe/H]
to intermediate-res RVS

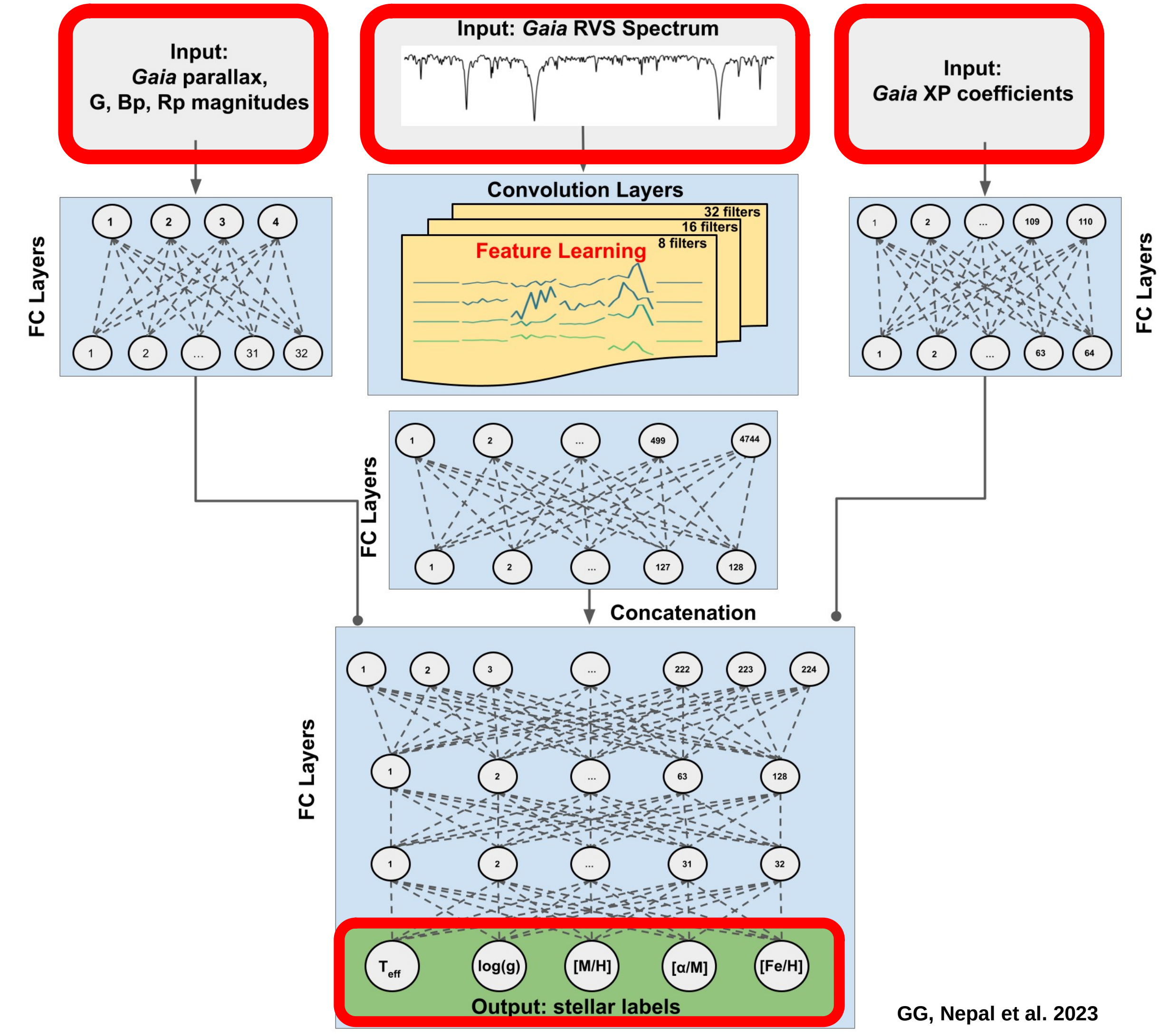
R~11400

Training sample



GG, Nepal et al. 2023

A hybrid Convolutional Neural-Network for *Gaia*-RVS analysis



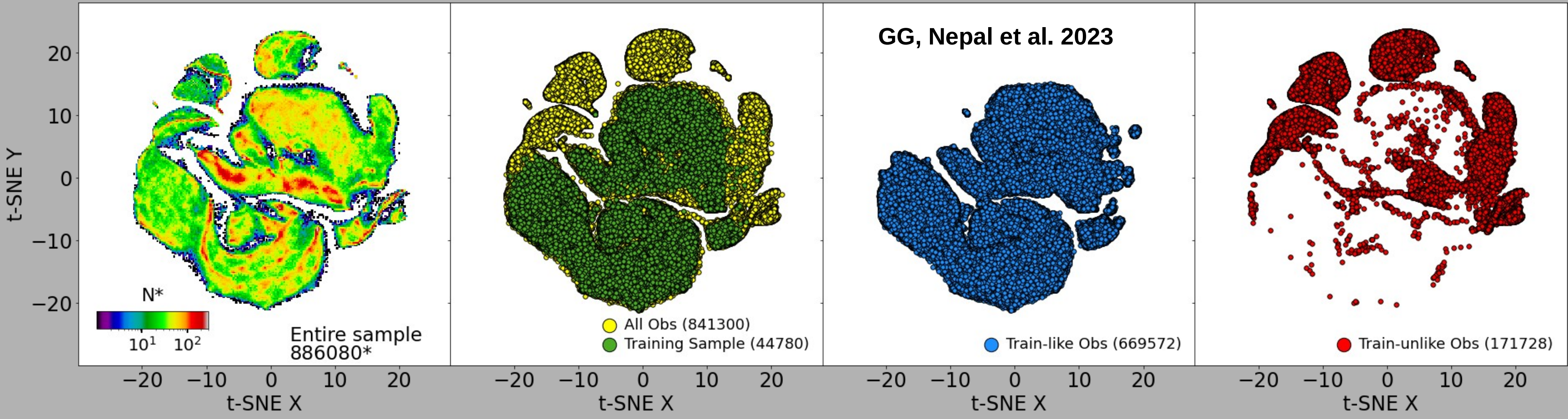
→ **Observed sample:**
841300 RVS stars

→ **Prediction time**
3300 stars / second

How to ensure that a label falls within the training sample limits ?

→ **Labels within T_{eff}, log(g), [M/H], [α/M], [Fe/H], G, and parallax limits of training sample.**

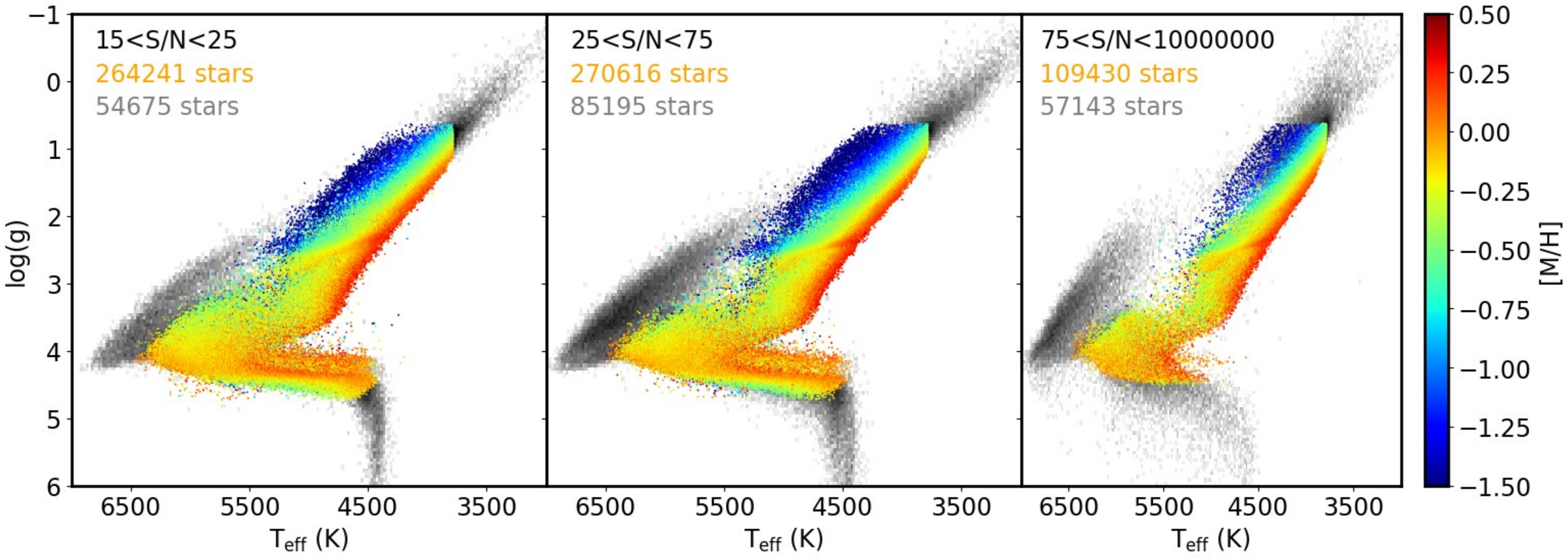
→ **t-SNE classification of RVS spectra**



→ **644287 RVS stars within TS**

Robust estimates of T_{eff}, log(g), [M/H] for 690000 *Gaia* stars

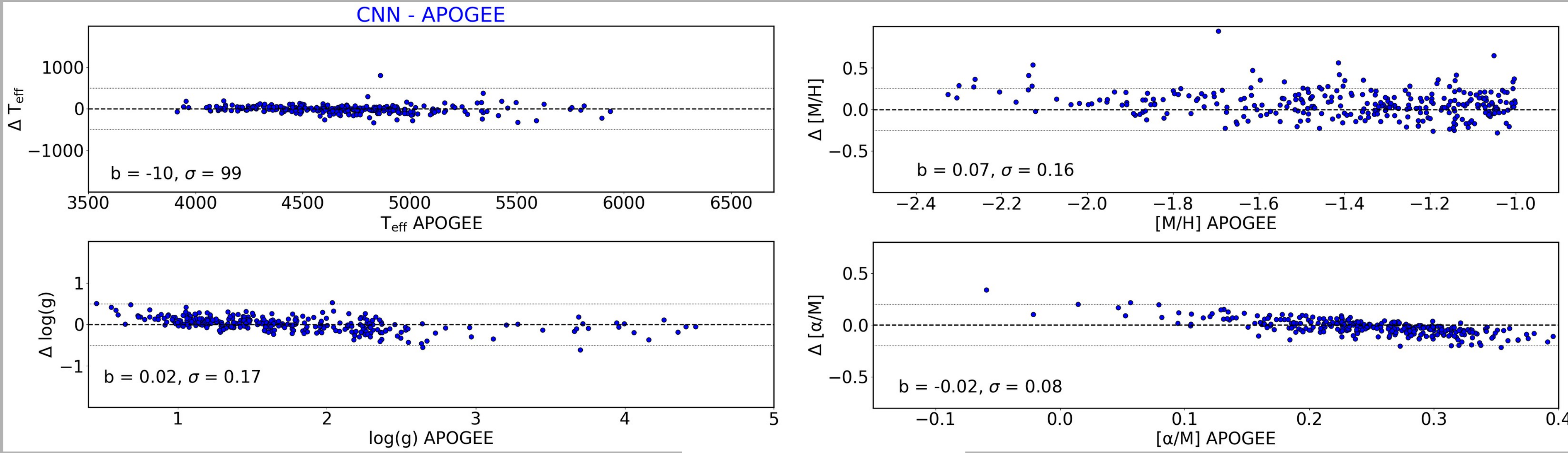
GG, Nepal et al. 2023



→ **By adding magnitudes, parallaxes and XP data, CNN is able to break spectral degeneracies in *Gaia* RVS spectra.**

→ **CNN results are as good as the training set can be.**

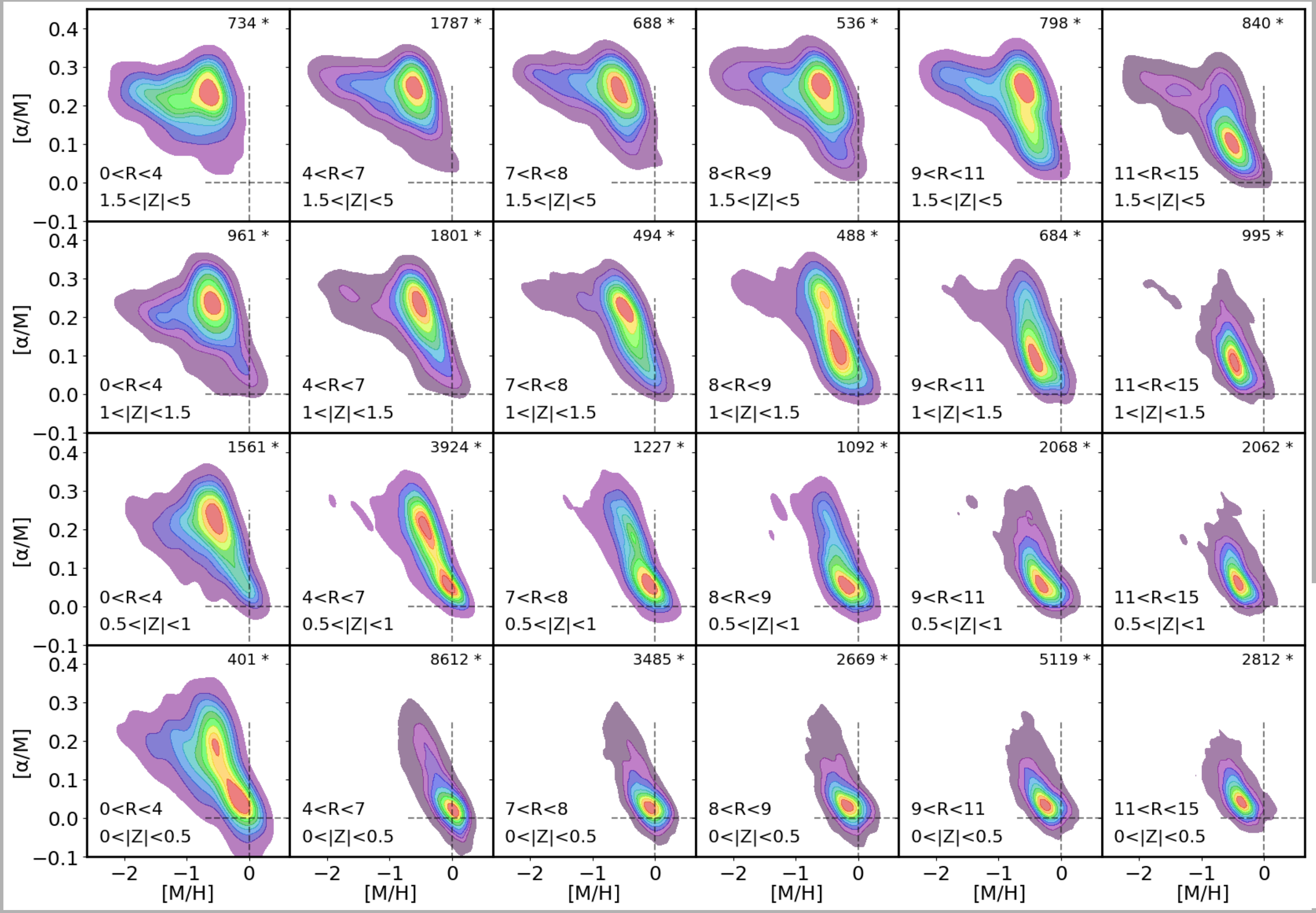
CNN performances for halo stars
→ 15<S/N<25



GG, Nepal et al. 2023

→ CNN provides precise and accurate labels down to [M/H]=-2.4 dex

Chemical cartography of the Milky Way, for Inner to Outer regions with *Gaia* and CNN

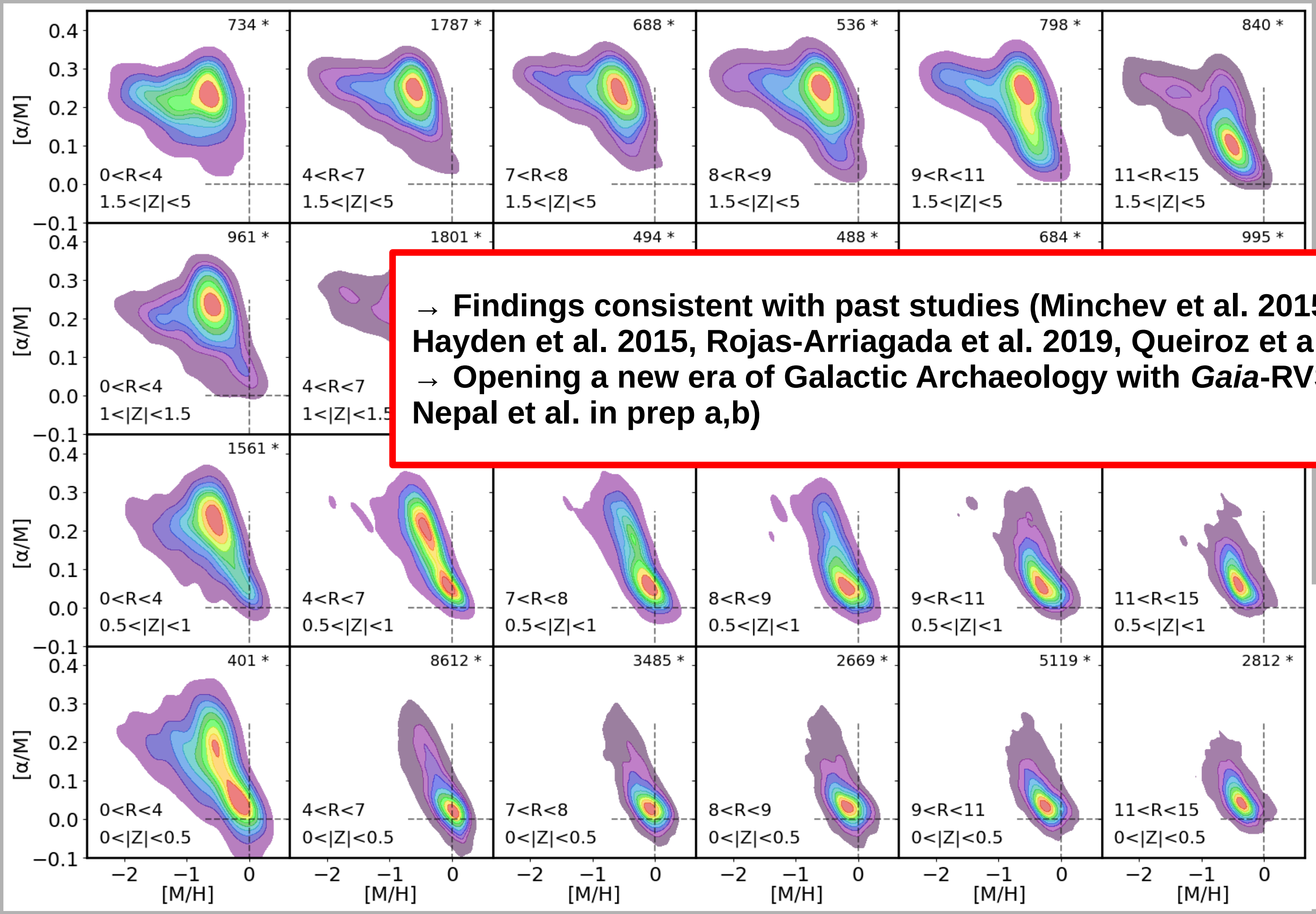


15<S/N<25

GG, Nepal et al. 2023

Chemical cartography of the Milky Way, for Inner to Outer regions with *Gaia* and CNN

15<S/N<25

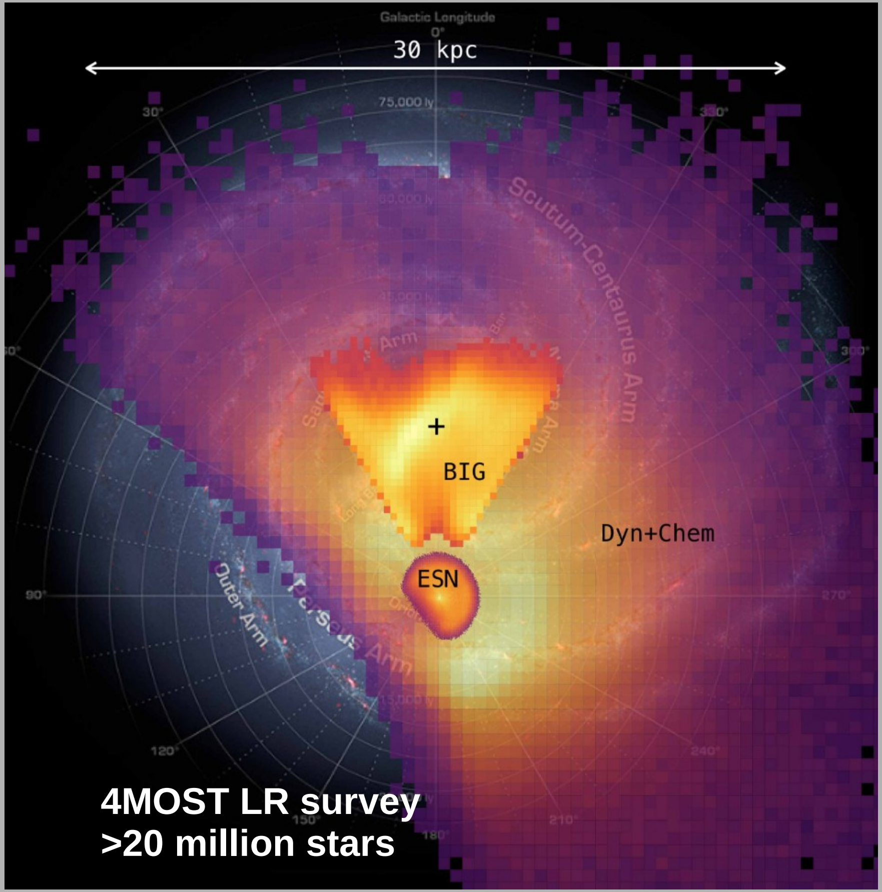


→ **Findings consistent with past studies (Minchev et al. 2015, Anders et al. 2014, Hayden et al. 2015, Rojas-Arriagada et al. 2019, Queiroz et al. 2020, 2021).**
→ **Opening a new era of Galactic Archaeology with *Gaia*-RVS (Guiglion et al. In prep, Nepal et al. in prep a,b)**

GG, Nepal et al. 2023

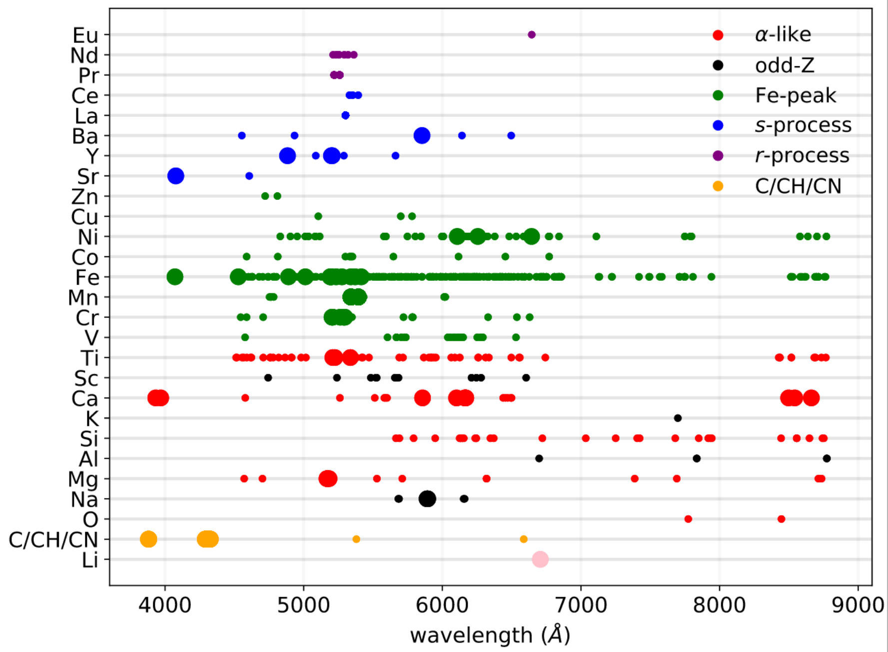
Why using CNN on low-res spectra ?

→ **4MIDABLE-LR Disc and Bulge surveys (Chiappini et al. 2019)**



4MIDABLE-LR ESO proposal 2020

>20 elements to be measured at R=5000



4MIDABLE-LR ESO proposal 2020

→ **Developing CNN for 4MIDABLE-LR D1(>) spectral analysis.**

Summary:

- Hybrid CNN is an optimal method for combining full Gaia data product
  → Leveraging the large set of low S/N RVS spectra

- CNN parametrization is fast and robust (several 10^3 stars per second)

Insights:

- Future spectroscopic surveys will strongly benefit from CNNs

- Standard spec. and ML methods complement each other

- CNN parametrization mainly reliable within the training sample limits
  → The training sample should be built in a pro-active way

guiglion@mpia.de

ZENTRUM FÜR ASTRONOMIE

DFG Deutsche Forschungsgemeinschaft German Research Foundation